DOCUMENT RESUME

ED 366 909 CS 011 586

AUTHOR Reinwein, Joachim
TITLE Etude du contexte visuel en lecture a l'aide d'une nouvelle technique d'auto-presentation segmentee (APS). (Study of Reading Context with the Aid of a New Segmented Auto-Presentation Technique [CAPS]).
PUB DATE Dec 93
NOTE 32p.
PUB TYPE Reports - Research/Technical (143)
LANGUAGE French

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Analysis of Variance; Cognitive Processes; Computer Assisted Instruction; Elementary Education; Error Analysis (Language); Foreign Countries; *Reading Comprehension; Reading Diagnosis; Reading Research; *Text Structure; Visual Stimuli
IDENTIFIERS Quebec (Montreal); *Text Processing (Reading); Zigzag Computer Program

ABSTRACT

This paper presents a new computer program called Zigzag, suited for investigating readers' mental activity during text processing. The program permits data collection of two different kinds, namely reading time and reading errors for each of the words read. The second dependent measure, reading errors, provides a better control of the readers' text comprehension and permits the restriction of statistical analyses of the time data to correctly identified words. Subjects of this study were French-speaking third and sixth graders living in the greater Montreal (Quebec, Canada) area. One major goal was to examine the effect of visually presented contextual information on text processing at two developmental levels and to examine the readers' dependence on visual information in an interactional perspective. The paper analyzes the results for the experimental text as a whole, but also with respect to some linguistic variables, such as grammatical category, word position in the sentence, or other variables. Results indicate major differences between the data and the time data obtained by means of other on-line techniques with respect to the same linguistic variables. The paper concludes with a more detailed discussion of these differences. Four tables and two figures of data are included. Contains 27 references. (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 366 909

STUDY OF READING CONTEXT WITH THE AID OF A NEW SEGMENTED

AUTO-PRESENTATION TECHNIQUE (CAPS)

Etude du contexte visuel en lecture à l'aide d'une nouvelle technique d'auto-présentation segmentée (APS)

Joachim Reinwein

Abstract

This article presents a new computer program called Zigzag (Ciesielski & Reinwein, 1989), suited for investigating the readers' mental activity during text processing. The program permits to collect data of two different kinds, namely reading time and reading errors for each of the words read. The second dependent measure, reading errors, provides a better control of the readers' text comprehension and permits to restrict the statistical analysis of the time data to correctly identified words.

The subjects of the present experiment were French speaking third- and sixth-graders living in the greater Montreal. One major goal of the experiment was to study the effect of visually presented contextual information on text processing at two developmental levels and to examine the readers' dependence on visual information in an interactional perspective (context x age). The experimental results are analyzed (by means of ANOVAs) for the experimental text as a whole, but also with respect to some linguistic variables, such as grammatical category, word position in the sentence, or other variables (ANOVAs and correlation analysis). There are some major differences between our own data and the time data obtained by means of other on-line techniques with respect to the same linguistic variables. These differences are discussed in more detail in the rest of the paper.

AS011556

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Reinwein

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

Etude du contexte visuel en lecture à l'aide d'une nouvelle technique d'auto-présentation segmentée (APS)

Joachim Reinwein
Département de linguistique
Université du Québec à Montréal

Décembre 1993

L'activité cognitive du lecteur, on le sait, n'est pas directement accessible à l'observation. C'est la raison pour laquelle l'étude de l'acte de lire doit s'appuyer sur des manifestations externes du lecteur telles que les mouvements oculaires, l'oralisation de la lecture, les réponses à un test de compréhension, etc.

De façon générale, on peut distinguer les techniques d'observation du lecteur selon qu'elles s'appuient sur des manifestations externes obtenues pendant ou après la lecture du texte. Les techniques d'auto-présentation segmentée (APS ou "on line") permettent d'observer le lecteur pendant la lecture. Le texte à lire y est présenté de sorte que le lecteur ne voit, à la fois, qu'une partie plus ou moins restreinte (cf. Pynte, 1983, par exemple). Chaque fois que le lecteur appuie sur un bouton, le segment suivant lui est présenté. Ainsi, la durée de présentation de chaque segment est sous le contrôle direct du lecteur. Le laps de temps écoulé entre deux commandes successives donne lieu à la mesure dépendante appelée temps de lecture (ou TEL = temps d'exposition en lecture; Zagar, 1988). On s'attend à ce que le lecteur n'appuie sur le bouton que lorsqu'il a "compris" le segment lu.

Les techniques APS connues ne disposent actuellement que d'une seule mesure, le temps. C'est à partir de l'indice de temps que l'acte de lire est étudié, la complexité du traitement cognitif du texte par le lecteur étant tributaire soit de la nature du texte (difficile ou facile), soit de l'aptitude du lecteur (bon lecteur ou lecteur faible). On suppose généralement que plus le temps de lecture est élevé, plus le texte (ou un segment de celui-ci) est difficile à lire ou encore, moins le lecteur est habile. Bien que cette hypothèse de travail se soit avérée extrêmement productive et que les études en temps réel constituent entre-temps un paradigme de recherche solidement établi, on constate cependant que, jusqu'ici, les techniques APS ont été confinées aux seules recherches expérimentales dites en laboratoire. Et pour cause: basé sur le seul indice de temps, il n'existe pas de moyen sûr pour vérifier si le lecteur comprend réellement ce qu'il lit. En situation d'évaluation scolaire par exemple, un lecteur, sachant que la vitesse servira de critère d'évaluation, peut très bien faire défiler les segments d'un texte rapidement et de façon purement mécanique, c'est-à-dire sans comprendre ce qu'il lit.

Ceci dit, nous croyons que les temps de lecture doivent être interprétés avec beaucoup de circonspection, même lorsqu'ils proviennent de lecteurs qu'on pourrait qualifier "de bonne foi" et ce, pour la raison suivante. Il est bien connu que la

compréhension d'un texte ne fonctionne pas selon le principe du tout ou rien: le même texte peut être compris de différentes manières et à différents degrés. Un lecteur est libre d'accorder au texte un traitement sémantique plus ou moins profond, le traitement sémantique pouvant varier entre une lecture superficielle (lecture "en diagonale", lecture "du premier degré", etc.) et une lecture plus profonde (lecture "entre les lignes", lecture critique, lecture herméneutique, etc.). Or, si l'on admet que le lecteur est libre de traiter un texte plus ou moins en profondeur, il faut également admettre que ces différences, dans le traitement du texte, puissent se traduire au niveau des ressources cognitives à investir dans l'activité de lecture et, par conséquent, au niveau du temps de lecture. Cette affirmation n'a rien d'original. Au fond, il ne s'agit ici que d'une variante de l'argument voulant que la complexité cognitive de l'acte de lire, attribuable au degré d'aptitude du lecteur ou au degré de difficulté du texte, se traduise par un temps de lecture plus ou moins élevé.

Pour pouvoir tenir compte de la profondeur du traitement des informations par les sujets expérimentaux, certaines études psychologiques disposent d'une deuxième mesure dépendante appelée "acuité". Cette mesure tient compte du nombre d'erreurs de compréhension en post-test. Cette manière de procéder permet de relativiser le facteur temps. Dans l'étude de Booher (1975) par exemple, un des cinq traitements expérimentaux produit à la fois le temps de lecture le plus court et le nombre d'erreurs le plus élevé. Il est évident que de s'appuyer, dans ces circonstances, uniquement sur l'indice de temps, fournirait un portrait incomplet, voire faussé, du comportement humain.

Pour pouvoir faire face à ce problème, la technique d'auto-présentation segmentée que nous avons conçue sous le nom de Zigzag (Ciesielski et Reinwein, 1989) dispose d'une deuxième mesure susceptible de fournir un indice de compréhension à la lumière duquel les temps de lecture peuvent être interprétés. Plus précisément, cette technique permet d'obtenir un temps de lecture pour chaque mot du texte, de façon à ce qu'un indice de type "réussite - échec" lui soit assigné. L'obtention de cette deuxième mesure est rendue possible grâce à la particularité de la technique Zigzag qui oblige le lecteur à choisir continuellement entre deux mots dont un seul est parfaitement acceptable dans le contexte des segments textuels déjà présentés et / ou à venir. Il s'agit là d'une tâche de décision lexicale entre deux mots existants: seul le mot cible permet de reconstruire le texte à lire, l'autre mot étant un mot distracteur sélectionné selon une procédure aléatoire (voir plus loin). Soulignons que la deuxième mesure permet de limiter le calcul statistique des temps de lecture aux

réussites, le choix du mot distracteur par le lecteur étant considéré comme indice d'une incompréhension locale du texte.

Etant donné la tâche expérimentale de la technique Zigzag, il est possible de faire varier ce que nous appelons l'ampleur du contexte linguistique visuel affiché à l'écran (bref: contexte visuel). En faisant varier l'ampleur du contexte visuel qui précède ou suit le mot cible, nous croyons pouvoir étudier l'impact que peut exercer la présence d'informations visuelles sur la tâche expérimentale.

Une mise en garde s'impose ici. L'ampleur du contexte visuel n'est pas à confondre avec l'ampleur, tout court, du segment linguistique affiché à l'écran cathodique. En effet, il est bien connu que l'ampleur du segment linguistique présenté à l'écran peut varier d'une technique d'auto-présentation à l'autre (Ferres et Aaronson, 1981: un seul mot à la fois; Mitchell et Green, 1978: segments de trois mots; Graesser, Hoffman et Clark, 1980 et Haberlandt, 1984 et 1988: segments de la longueur d'une proposition ou phrase complexe; etc.). Néanmoins, quelle que soit, dans ces études, l'ampleur du segment affiché, il reste que ces variantes reposent sur la mise en correspondance entre le temps et le segment visuel *dans son ensemble* (et non une partie de celui-ci). Autrement dit, dans le cas de ces variantes, la notion de *contexte* est ramenée à celle de *contexte non visuel.*

Bien que la technique Zigzag permette de faire varier aussi bien le contexte visuel précédent que le contexte visuel suivant le mot cible, on ne fera varier, dans cette étude, que le dernier.

De façon générale, on peut présumer que l'utilité du contexte visuel sur le traitement des informations par le lecteur varie d'un mot à l'autre, selon des caractéristiques intrinsèques et de la place qu'il occupe au sein de la phrase et du texte. Pour mieux cerner l'effet potentiel du contexte visuel sur le traitement cognitif des mots, la variable Contexte sera étudiée en lien avec différentes variables linguistiques et textuelles qui, en partie, ont déjà fait l'objet d'analyses en temps réel:

Le nombre de lettres: Généralement, le temps de lecture est positivement corrélé avec le nombre de lettres du mot (Mitchell et Green, 1978; Just, Carpenter et Woolley, 1982; Haberlandt et Graesser, 1985). D'après une expérience de O'Regan (1986, non publié) et dont une partie des résultats est présentée dans Levy-Schoen (1988), l'effet de longueur des mots n'est pas réductible à leur fréquence: l'effet de longueur

est significatif même si la fréquence des mots est contrôlée expérimentalement, et il se manifeste aussi bien pour les mots fréquents que les mots rares.

La catégorie grammaticale (cf. Aaronson et Ferres, 1983, 1984a, 1984b).

La position des mots dans la phrase: Un des phénomènes les plus stables parmi les différentes études d'observation en direct, est le temps de lecture prolongé en fin de phrase (Mitchell et Green, 1978; Green, Mitchell et Hammond, 1981; Just, Carpenter et Woolley, 1982; Haberlandt et Graesser, 1985; Zagar, 1988); ceci est interprété généralement comme indice d'une charge cognitive élevée qui traduit une activité d'intégration des mots et syntagmes de la phrase.

La position des mots dans le texte: L'effet de la position sérielle des mots ou phrases dans le texte a été démontré à plusieurs reprises (Mitchell et Green, 1978; Cirilo et Foss, 1980; Haberlandt, Berian et Sandson, 1980; Graesser, Hoffman et Clark, 1980; Haberlandt, 1984; Frochot, Zagar et Fayol, 1987).

Finalement, l'effet du contexte visuel sur le temps de lecture du syntagme nominal et ses parties constitutives, sera également étudié.

L'étude s'inscrit dans la lignée des travaux qui questionnent le bien-fondé des hypothèses de Just et Carpenter (1980). D'après ces auteurs, le traitement cognitif de chaque mot se fait au moment même de sa lecture ("immediacy assumption"), et l'oeil repose sur le mot tout le temps nécessaire à son traitement ("eye-mind assumption"). Just et Carpenter (1984) ont par la suite restreint le domaine d'application de leurs hypothèses au traitement *lexical* des mots (voir aussi Ehrlich et Rayner, 1983 et Mitchell, 1984). Pour ce qui est des techniques APS, Mitchell et Holmes (1985) n'excluent pas la possibilité d'un décalage perceptivo-cognitif de sorte que les lecteurs fassent avancer le segment de texte affiché à l'écran même s'ils n'ont pas encore entièrement résolu la tâche cognitive qui s'y rapporte ("processing spillover"). Dans l'ensemble, cette discussion fait bien voir les limites interprétatives auxquelles sont assujetties les études dont la seule mesure dépendante est le temps.

La présente étude porte sur l'apprenti-lecteur à l'école primaire. L'ajout d'une variable Niveau scolaire (3e vs 6e année du primaire) permettra de situer les phénomènes analysés dans une perspective développementale. Les deux niveaux scolaires retenus correspondent sur le plan administratif à la fin des premier et

deuxième cycles de l'école primaire au Québec et marquent deux niveaux d'habiletés nettement distincts en lecture.

Description de l'outil Zigzag

La première version de l'outil informatisé Zigzag (bref: ZZ) a été utilisée expérimentalement dès 1988 (Ciesielski et Reinwein, 1989). Depuis, plusieurs modifications ont été apportées à l'outil. Les logiciels Générateur ZZ 3.0 (16K) et Test ZZ 4.2 (76K), écrits en LightSpeed Pascal 2.11 de Think Technologies, représentent la version utilisée dans cette étude. L'outil Zigzag est utilisable avec des appareils de type Macintosh. Le nom de l'outil s'est inspiré du mouvement manuel que l'on exécute lors de l'utilisation du logiciel avec la souris. Dans la version actuelle, l'utilisateur peut en principe se servir de la souris ou du clavier; dans la présente expérimentation, tous les lecteurs devaient cependant utiliser le clavier.

Le logiciel Générateur ZZ, qui peut opérer sur tout texte rédigé avec un logiciel de traitement de texte, permet de générer un texte selon certains paramètres. Le logiciel Test ZZ produit l'interface correspondant à ces paramètres. Lors du test, le logiciel Test ZZ enregistre le temps de lecture pour chaque mot du texte et indique s'il y a réussite ou non.

La technique Zigzag présente le texte à lire de façon à ce que le lecteur doive continuellement choisir entre deux mots, le mot cible et le mot distracteur. Cette technique consiste en une sorte de tâche de décision lexicale entre deux mots existants, mais dont seulement l'un - le mot cible - permet de reconstruire le texte à lire. Pour sélectionner le mot cible, le lecteur appuie le plus vite possible sur l'une des deux touches préalablement définies sur le clavier. Dès que le lecteur a identifié le mot cible, le texte avance d'un mot. En voici un exemple.

Figure 1
Séquence de cinq écrans présentés successivement au lecteur à l'aide du logiciel Test ZZ

insérer ici figure 1

La figure 1 présente cinq écrans que le lecteur voit successivement. Dans chacune des colonnes à deux rangées, figure un (1) mot, le mot cible, qui permet au lecteur de

reconstruire le texte "caché". La zone activée de l'écran étant toujours la colonne la plus à gauche, le lecteur essaie de sélectionner le plus rapidement possible, à l'aide du clavier ou de la souris, le premier mot du texte, i.e. le mot cible **Mon**. Dès que le lecteur l'a sélectionné, l'écran 1 est remplacé par l'écran 2. Le lecteur a alors l'impression que les colonnes affichées défilent vers la gauche. En même temps, le mot cible sélectionné est affiché dans la fenêtre du haut. Si le lecteur choisit la mauvaise réponse (dans ce cas-ci, le mot distracteur **les**), l'écran 1 reste inerte aussi longtemps que le lecteur ne change pas d'idée pour choisir le mot cible. Tous les essais infructueux sont enregistrés et indiquent le nombre d'erreurs commises par le lecteur. Le laps de temps qui s'écoule entre deux écrans successifs est enregistré, en millisecondes (intervalle minimal: 16.7 ms), et constitue la principale mesure dépendante du test. Dès que le lecteur a choisi le deuxième mot cible, le troisième écran s'affiche, et les deux premiers mots cibles du texte sont présentés dans la fenêtre du haut, comme contexte précédent. Le contexte précédent étant l'un des paramètres contrôlables de l'outil Zigzag, les mots cibles, une fois sélectionnés par le lecteur, sont affichés dans cette fenêtre qui était vide au début. La figure 1 montre cinq écrans successifs par lesquels le lecteur identifie les cinq premiers mots du texte (**Mon père, ma mère, mon ...**).

Selon un premier principe régissant l'outil Zigzag, les mots cibles (= texte original) sont présentés aléatoirement, soit dans la première, soit dans la deuxième rangée de chaque colonne. L'emplacement aléatoire des mots cibles est une condition cruciale pour que le lecteur ne puisse anticiper les choix à faire sur une base mécanique. Prenons comme exemple le premier mot du texte, le mot cible **Mon**. Calculé sur le total des 150 générations individuelles du texte expérimental utilisé, le mot cible **Mon** figure à peu près le même nombre de fois en haut et en bas (73 et 77 fois, respectivement).

Selon un deuxième principe, le Générateur ZZ sélectionne aléatoirement les mots distracteurs dans l'ensemble des mots du texte original, en excluant toutefois la possibilité que, dans une colonne donnée, le mot distracteur soit identique au mot cible. Ce principe permet, lors de l'analyse des résultats, d'identifier l'impact relatif du mot cible et du mot distracteur sur les deux mesures dépendantes. Etant donné que la complexité cognitive de chaque sélection dépend aussi bien du mot distracteur que du mot cible, cette technique de contrebalancement aléatoire est indispensable. Pour donner un exemple: L'identification du premier mot cible de notre texte expérimental (**Mon**) est probablement plus facile en présence d'un mot distracteur

qui commence par une minuscule ou qui appartient à une autre catégorie grammaticale que celle du mot cible. Il est donc crucial, faute de pouvoir éliminer l'effet attribuable au mot distracteur spécifique, au moins de l'atténuer en générant autant de tests différents qu'il y a de lecteurs. Dans la présente étude, sur un total de 150 générations individuelles du texte expérimental utilisé, le premier mot cible **Mon** a été combiné avec 69 mots distracteurs différents (sur un total théoriquement possible de 81 mots distracteurs issus du même texte). C'est en effet grâce à cette variation des mots distracteurs que les mots d'un texte peuvent être comparés statistiquement.

Résumons: Le mot cible se distingue du mot distracteur non pas de façon absolue - le même mot peut être un mot cible ou un mot distracteur - mais en fonction du rôle qu'il joue dans la reconstruction du texte "caché": seuls les mots cibles permettent la reconstruction du texte original. Les contraintes syntaxiques et sémantiques exercées par le contexte linguistique (visuel et non-visuel) sur le choix du mot cible sont telles que, malgré la sélection aléatoire du mot distracteur, les cas où le mot distracteur serait acceptable au même titre que le mot cible sont rares au point d'être négligeables.

Méthode

Sujets. L'expérimentation a été menée à Montréal auprès de 300 enfants de langue française et issus de milieu socio-économique faible ou moyen. La moitié des enfants était en troisième année du primaire et l'autre moitié, en sixième. Au sein de chacun des deux niveaux scolaires, les enfants ont été assignés aléatoirement et en nombre égal aux cinq traitements expérimentaux (Contexte 0 à 4), c'est-à-dire à dix groupes de 30 enfants.

Matériel. Le texte expérimental intitulé "La pêche", d'une longueur de 123 mots, était extrait d'un livre de lecture s'adressant à des enfants de 3e année du primaire (Ciesielski et Reinwein: Les apprentis: lecture 3. Montréal: Etudes Vivantes, 1985, p. 163). A partir de ce texte, 150 tests individuels ont été générés à l'aide du logiciel Générateur ZZ, chaque test étant destiné à un lecteur de troisième année et un lecteur de sixième année.

Lors de l'expérimentation, les 150 tests ont été présentés à l'aide du logiciel Test ZZ en cinq versions expérimentales, trente tests par version (Contexte 0 à 4). La figure 2 illustre les versions expérimentales qui permettent aux lecteurs, à des degrés variables, de se servir du contexte visuel qui suit pour identifier le mot cible **Mon**:

Figure 2
Exemple des cinq versions expérimentales (Contexte 0 à 4)

insérer ici figure 2

Les lecteurs de la version "Contexte 0" ne disposaient pas d'indices contextuels pouvant les aider. Dans les autres versions par contre (Contexte 1 à 4), les lecteurs avaient accès, à des degrés variables, à des indices contextuels.

Pour communiquer avec l'ordinateur, les lecteurs devaient se servir de deux touches du clavier, en l'occurrence les touches "4" et "1" du clavier numérique. Notons que ces touches sont différentes des touches "4" et "1" du clavier principal. Situées à l'extrémité droite du clavier, elles sont placées l'une au-dessus de l'autre et donc de façon analogue à leur affichage à l'écran.

Procédure. Les versions expérimentales et le texte servant de pratique ont été générés avant l'expérimentation proprement dite. Le test s'est fait individuellement pour chaque enfant dans une salle de classe aménagée à cet effet. Avant l'expérimentation proprement dite, les lecteurs ont pu se familiariser avec l'ordinateur et, plus spécifiquement, avec l'outil Zigzag. Ainsi, chaque enfant a pu s'entraîner à l'aide d'un paragraphe d'une trentaine de mots.

Compilation des données. Les deux mesures dépendantes de l'expérimentation étaient le temps de lecture et le nombre d'erreurs. Pour l'analyse des temps de lecture, seuls les temps correspondant aux mots réussis ont été retenus. Cette manière de procéder augmente la fiabilité des temps de lecture comme indice des processus cognitifs sous-jacents. En écartant de la base de données les temps de lecture correspondant aux erreurs, on s'assure de ne retenir que les données les plus systématiques et donc révélatrices pour les processus analysés, la variabilité des temps de lecture des erreurs étant supérieure à celle des réussites. D'autre part, on évite un biais systématique dû à la façon de mesurer les temps de lecture. En effet, en cas d'erreur (i.e. lors de la sélection d'un mot distracteur), l'écran reste figé,

ce qui oblige le lecteur à appuyer sur l'autre touche de réponse. Or, ce n'est qu'au moment où le lecteur appuie sur la touche juste - action qui déclenche le changement d'écran - que se termine le délai correspondant à l'erreur en question. Ceci se traduit bien sûr par un délai supplémentaire qui est peu révélateur des processus de lecture à analyser. Cependant, même si les mots erronés sont exclus du calcul des temps de lecture, ils donnent lieu à une deuxième mesure fort éloquente, le nombre d'erreurs commises.

Devant le problème des temps associés aux réponses erronées, il y avait deux choix méthodologiques, à savoir remplacer le temps d'une réponse erronée par un temps fictif nettement plus élevé que le temps normal, ou le déclarer comme "valeur manquante". Le choix d'un temps élevé introduirait un biais semblable à celui décrit ci-haut. Ce biais se produirait même dans l'éventualité que le lien entre le temps moyen de réponse chez un enfant et son nombre d'erreurs serait strictement linéaire. Nous avons donc choisi de considérer les temps associés aux erreurs comme des "valeurs manquantes". Cette décision entraîne cependant un autre problème, celui d'avoir très peu d'enfants ayant répondu à tous les mots du texte. Il n'a donc pas été possible d'avoir une banque de données assez complète pour étudier les effets des variables expérimentales et de contrôle en tenant compte de l'effet sujet. Aussi, les calculs statistiques ont été effectués sur deux bases de données.

Pour étudier avec le maximum de souplesse les effets de la version, du niveau scolaire et d'autres variables de contrôle sur les temps de réponse et le nombre d'erreurs aux 123 mots, la première base de données comptait 36900 observations (123 mots X 300 enfants) et les variables suivantes: l'identification de l'enfant, le niveau scolaire, le contexte, la position du mot dans la phrase, le syntagme et la position de la phrase dans le texte.

L'autre base de données, plus classique, rendait compte des temps de réponse aux 123 mots, en cas de réussite, pour chacun des enfants. Cette banque contenait 300 observations (150 enfants X deux niveaux scolaires) et 136 variables: les 123 mots du texte, le code d'identification de l'enfant, son niveau scolaire et le contexte. Cette matrice a permis d'étudier les effets du niveau et du contexte sur le nombre d'erreurs de chaque enfant et le temps moyen de réponse de chaque enfant pour l'ensemble des mots ou certains mots sélectionnés. Le nombre d'erreurs de chaque enfant correspondait au nombre de temps manquants pour l'ensemble des mots. Le temps moyen de réponse d'un enfant était calculé en faisant la somme de ses temps de

réponses pour l'ensemble des mots et en divisant cette somme par le nombre de mots auxquels il donnait une réponse correcte.

Pour étudier les effets des variables sur le temps de réponse, des analyses de variance ont été effectuées, suivies au besoin de tests du LSD. Par ailleurs, les liens entre le nombre d'erreurs et le temps moyen de réponse de chaque enfant ont été mesurés par la corrélation de Pearson.

Résultats

L'effet du contexte visuel sur les mots en général

Pour l'ensemble des mots du texte, deux analyses de variance ont été calculées, une par mesure (temps, nombre d'erreurs). Chaque analyse de variance comprenait deux variables inter-sujets, à savoir le contexte (cinq valeurs) et le niveau scolaire (deux valeurs). Notons toutefois que, dans plusieurs des analyses de variance subséquentes, la variable Contexte a été réduite à deux niveaux, opposant ainsi le traitement Contexte-0 (absence de tout contexte) aux traitements Contexte-1 à 4 (présence d'un contexte).

Pour ce qui est du temps de lecture, les variables Contexte [$F(4,32681) = 5.16$, $p < 0,0001$] et Niveau [$F(1,32681) = 1874,24$; $p < 0,0001$] de même que l'interaction Contexte X Niveau [$F(4,32681) = 3,69$; $p < 0,01$] sont significatives. Au tableau 1, les temps de lecture sont présentés séparément pour les deux niveaux scolaires.

Tableau 1
Temps de lecture moyen (en ms) par contexte et niveau scolaire

	3e année	6e année
Contexte:		
0	3109	2230
1	3220	2148
2	3254	2342
3	3273	2213
4	3112	2223

*Test du LSD: écart minimal significatif = 0,117 (en 3e) et 0,076 (en 6e), $p < 0.05$

D'après le test du LSD, en 3e année, les groupes aux deux extrêmes, Contexte-0 et Contexte-4, prennent significativement moins de temps que les trois groupes intermédiaires Contexte-2, Contexte-3 ($p < 0,05$) et Contexte-1 ($p < 0,06$). En 6e année, les groupes aux deux extrêmes et le groupe Contexte-3, sans se distinguer l'un de l'autre, prennent significativement moins de temps que le groupe intermédiaire Contexte-2. Ainsi, en faisant abstraction du groupe Contexte-1 en 6e année, on constate que les groupes situés aux deux extrêmes prennent moins de temps que le ou les groupes intermédiaires.

Pour ce qui est du nombre d'erreurs, l'effet des variables Contexte [$F(4,290) = 3,04$, $p < 0,05$] et Niveau [$F(1,290) = 99,16$, $p < 0,0001$] s'avère significatif. Les versions Contexte-0 (15,53 %), Contexte-2 (15,63 %), Contexte-3 (13,50 %) et Contexte-4 (13,78 %) ne se distinguent pas significativement entre elles (LSD: écart minimal significatif = 2,6; $p < 0,05$), la seule exception étant la version Contexte-1 dont le pourcentage d'erreurs est de 11,70 % seulement.

L'effet attribuable à la catégorie grammaticale et la longueur des mots

Les facteurs Catégorie grammaticale et Longeur des mots ont été étudiés au sein d'une analyse de variance à quatre facteurs (contexte, niveau, catégorie, longueur). Dans cette analyse de variance calculée sur le temps de lecture, le facteur Contexte a été réduit à deux niveaux (avec et sans contexte) et le facteur Longueur, à trois niveaux (mots de une ou deux lettres, mots de trois à cinq lettres, mots de plus de cinq lettres). L'analyse de variance a permis de déceler deux effets d'interaction et un effet d'interaction près du seuil de signification, tous en lien avec le niveau scolaire [Niveau X Catégorie: $F(9,32654) = 2,38$; $p < 0,05$; Niveau X Longueur: $F(2,32654) = 3,21$; $p < 0,05$; Niveau X Contexte: $F(1,32654) = 3,24$; $p < 0,01$].

Les analyses subséquentes ont donc été effectuées pour les deux niveaux scolaires séparément. Il en ressort que les facteurs Catégorie grammaticale [3e année: $F(9,15687) = 32,86$; $p < 0,0001$; 6e année: $F(9,15687) = 80,70$; $p < 0,0001$] et Longueur [3e année: $F(2,15687) = 6,16$; $p < 0,01$; 6e année: $F(2,15687) = 47,36$; $p < 0,0001$] étaient significatifs aux deux niveaux scolaires, de même que l'interaction Catégorie x Longueur [3e année: $F(11,15687) = 20,29$; $p < 0,0001$; 6e année: $F(11,15687) = 37,56$; $p < 0,0001$].

Pour ce qui est du facteur Contexte, les lecteurs des deux niveaux scolaires se sont comportés différemment. Cette variable était significative en 3e année, mais non en 6e année [3e année: $F(1,15687) = 5,05$; $p < 0,05$; 6e année: $F(1,15687) < 1$]. D'après le test de LSD, les lecteurs de 3e année lisaient les mots significativement moins vite en contexte (avec contexte = 3215 ms, sans contexte = 3109 ms; écart minimal significatif = 92 ms., $p < 0,05$), contrairement aux lecteurs de 6e année (avec contexte = 2232 ms, sans contexte = 2230 ms).

Une interaction analogue s'est dégagée pour la deuxième mesure dépendante, soit le nombre d'errreurs. Seuls les lecteurs de 3e année faisaient significativement plus d'erreurs sans contexte (3e année: 16,48% vs 14,36%, $Chi^2 = 10,47$, 1 ddl, $p < 0,001$; 6e année: 8,78% vs 7,84%, $Chi^2 = 3,55$, 1 ddl, ns). La comparaison des deux mesures dépendantes chez les lecteurs de 3e année suggère que l'absence du contexte est liée à l'augmentation du nombre d'erreurs.

Le tableau 2 présente, pour chaque mot du texte, les deux mesures en fonction de la catégorie grammaticale:

Tableau 2

Temps de lecture moyen (en ms) et nombre d'erreurs moyen (en %) calculés par catégorie grammaticale et par niveau scolaire

	3e année		6e année	
	Temps* (en ms)	Erreurs** (en %)	Temps* (en ms)	Erreurs** (en %)
Pronom	3894	15,27	2848	9.03
Adjectif	3656	15,26	2556	9.04
Conjonction	3507	17,81	2766	11,14
Autre	3276	17,22	2269	9,11
Article	3216	18,94	2328	10,89
Préposition	3189	17,19	2381	11,04
Adverbe	3153	18,67	2281	9,39
Verbe	3079	10,78	2014	5,83
Auxiliaire	2998	15,47	1899	7,20
Nom	2837	10,36	1891	4,80

* Test du LSD calculé avec la variance résiduelle de l'ANOVA
Ecart minimal significatif: 3e année = 0.176, $p < 0.05$; 6e année= 0.112, $p < 0.05$
** Chi^2 = 138,10 (3e année) et 160,85 (6e année), 9 ddl, $p < 0,001$

Le tableau 2 montre que les soi-disant mots-outils (conjonction, préposition, article, pronom) nécessitent beaucoup plus de temps et donnent lieu à un nombre supérieur d'erreurs, comparativement aux mots pleins, en particulier les noms et les verbes.

Ce résultat fait par ailleurs comprendre l'effet d'interaction Catégorie X Longueur mentionné précédemment, les mots-outils étant généralement des mots plus courts que les mots pleins. Aux deux niveaux scolaires, les mots nécessitent d'autant plus de temps qu'ils sont courts (mots courts: moy = 3412 ms en 3e année et 2469 ms en 6e année; mots intermédiaires: moy = 3223 ms et 2335 ms; mots longs: moy = 2973 ms et 1912 ms; différence minimale significative en 3e et 6e = 89 ms et 57 ms, respectivement).

Quant à la deuxième mesure dépendante, on constate que les mots courts engendrent plus d'erreurs (en 3e: 18,00 %; en 6e: 10,46 %) que les mots intermédiaires (en 3e: 15,04 %; en 6e: 8,22 %), et ces derniers plus d'erreurs que les

mots longs (en 3e: 11,46 %; en 6e: 5,51 %; Chi^2 = 102,17 (3e année) et 99,96 (6e année), $p < 0,001$).

L'effet du contexte visuel sur le syntagme nominal

Dans cette analyse de variance à trois facteurs (contexte, niveau, syntagme), seuls les mots faisant partie d'un syntagme nominal de la structure "ARTICLE + NOM" ont été retenus (syntagme = facteur à deux niveaux, i.e. article et nom). L'analyse devait indiquer si la présence visuelle du nom facilite le choix de l'article qui le précède.

En ce qui concerne le temps de lecture, les facteurs Syntagme [$F(1,11291) = 325,30$; $p < 0,0001$] et Niveau [$F(1,11291) = 658,94$; $p < 0,0001$] était significatifs, mais non le facteur Contexte, ni l'interaction Syntagme x Contexte [$F(1,11291) < 1$]. Le temps de lecture de l'article (M = 2,924 s) était significativement plus long que celui du nom (M = 2300 ms; test du LSD = 69 , $p < 0,05$).

Quant à la deuxième mesure dépendante, l'article du syntagme engendrait significativement plus d'erreurs que le nom (Chi^2 = 13, 64, 1 dl, $p < 0,001$). Le tableau 3 révèle par ailleurs une interaction complexe entre le syntagme, le niveau et le contexte (cf. Discussion).

Tableau 3
Pourcentage d'erreurs du syntagme nominal (ARTICLE + NOM)
selon le contexte et le niveau scolaire

	Article		Nom	
	avec	sans	avec	sans
3e année	16,79	21,59	8,93	8,57
6e année	10,63	10,95	3,81	5,08

L'effet attribuable à la position des mots dans la phrase

L'effet attribuable à la position des mots dans la phrase a été étudié à l'aide d'une analyse de variance à trois facteurs (position, contexte, niveau, position), le facteur Position ayant trois niveaux (dernier mot, premier mot, autres mots). Dans cette analyse, le facteur Position [$F(2,32681) = 313,21$; $p < 0,0001$] avait un effet hautement significatif sur le temps de lecture (premier mot = 3701 ms, dernier mot = 2321 ms,

autres mots = 2653 ms; test du LSD: écart minimal significatif = 97). Notons aussi que l'effet du contexte [F(1,32681) = 3,43; p = 0,06] et l'interaction Contexte X Position [F(2,32681) = 2,43; p = 0,09] étaient près du seuil de signification. Comme le montre le tableau 4, cette interaction faible pourrait signifier que le premier mot d'une phrase est davantage affecté par la présence ou l'absence du contexte que le sont les autres mots de la phrase.

Tableau 4
Temps de lecture (en ms) et pourcentage d'erreurs
selon la position du mot dans la phrase et le contexte

	Avec contexte		Sans contexte	
	Temps	Erreurs	Temps	Erreurs
Dernier mot	2305	7,18	2387	7,96
Premier mot	3748	13,02	3506	16,04
Autres	2663	11,29	2612	12,77

De façon générale, les temps de lecture les plus longs correspondent aux pourcentages d'erreurs les plus élevés et les temps les pius courts, aux pourcentages les moins élevés.

L'effet attribuable à la position des mots dans le texte

Une dernière analyse de variance a été calculée pour les facteurs Contexte, Niveau et PositionTexte, ce dernier spécifiant la position du mot dans le texte (début, milieu, fin). Pour le temps de lecture, on constate un effet simple pour les facteurs Niveau [F(1,32411) = 1142,46; p < 0,0001] et Contexte [F(1,32411) = 4,32; p < 0,05] de même qu'un effet d'interaction Niveau X PositionTexte [F(2,32411) = 30,12; p < 0,0001]. Ces résultats signifient que les mots nécessitent plus de temps en contexte (avec contexte: 2709 ms, sans contexte: = 2652 ms; test du LSD: écart minimal significatif = 55 ms, p < 0,05). D'autre part, le temps de lecture diminue au fur et à mesure que le lecteur avance dans le texte (test du LSD, 3e et 6e années: premier tiers = 3031 ms, deuxième tiers = 2663 ms, troisième tiers = 2444 ms). L'effet d'interaction, quant à lui, est attribuable au fait que cette diminution de temps est plus prononcée chez les lecteurs de 3e année.

Le pourcentage d'erreurs augmente augmente au fur et à mesure que le lecteur avance dans le texte, mais de manière plus prononcée au début (Chi^2 = 49,62, 2 ddl, $p < 0,001$; premier tiers = 9,68 %, deuxième tiers = 12,06 %, troisième tiers = 12,32 %). Cette tendance générale se manifeste en 3e et en 6e années, avec et sans contexte, mais à des degrés variables. L'effet le moins prononcé se remarque en 6e année en absence du contexte, où les différences entre les tiers cessent d'être significatives (Chi^2 = 2,04, 2 ddl, $p > 0,05$).

Corrélations temps-erreurs

Pour examiner le rapport entre les deux mesures dépendantes, temps et erreurs, le coefficient de corrélation de Pearson a été calculé, d'une part, pour chacun des 123 mots du texte et, d'autre part, pour chacun des 300 enfants de l'expérience. Rappelons que le temps de lecture n'a été calculé que pour les mots réussis. Pour les mots du texte, le coefficient était significativement différent de l'hypothèse nulle aux deux niveaux scolaires [3e année: $r = 0,43$, $p < 0,001$; 6e année: $r = 0,62$; $p < 0,0001$]. Donc, pour un mot donnée, plus le nombre d'erreurs augmente, plus le temps de lecture est long pour ceux qui ont réussi à identifier correctement ce mot.

Il est à noter qu'on n'arrive pas au même constat lorsqu'on calcule la corrélation en fonction des sujets. Dans ce cas, le coefficient n'est différent de l'hypothèse nulle que pour les lecteurs de 6e année [3e année: $r = 0,04$; ns; 6e année: $r = -0,30$; $p < 0,001$]. La corrélation négative pour les lecteurs de 6e année signifie que plus un lecteur commet d'erreurs, plus son temps de lecture moyen pour les mots réussis est faible.

Discussion

L'effet du contexte visuel qui suit

L'objectif premier de l'étude était l'analyse, par le biais d'une technique d'auto-présentation segmentée appelée Zigzag, de l'effet du contexte *visuel* sur des lecteurs du primaire, à deux niveaux d'habiletés en lecture nettement distincts (3e vs.

6e années). Or, si l'étude montre qu'un tel effet existe réellement, elle montre également qu'il s'agit d'un effet plutôt fragile dans le cas des deux mesures dépendantes, le temps et le nombre d'erreurs. Il n'est pas exclu que ce fait soit en partie attribuable, malgré l'assignation au hasard des sujets aux groupes expérimentaux, à la non-équivalence des groupes (en l'occurrence, le groupe Contexte-1 en 6e année). Malheureusement, cette interrogation n'a pu être vérifiée dans cette étude, faute d'un prétest qui aurait pu servir de covariable à l'analyse de variance des données (cf. Reinwein, 1992). Pour cette raison, nous avons d'ailleurs préféré nous limiter, dans les analyses subséquentes, à une analyse dichotomique du facteur Contexte (absence vs présence du contexte), bien que cette simplification puisse être discutable.

Dans l'ensemble, l'étude montre que l'absence du contexte visuel n'entraîne pas de prolongation du temps de lecture. Abstraction faite du groupe "Contexte-1 de 6e année, c'est plutôt le contraire qui semble se produire: les temps de lecture, en l'absence du contexte, sont soit égaux à ceux en présence du contexte, soit inférieurs à ceux-ci. Quant au nombre d'erreurs, l'étude ne montre pas non plus une tendance nette, bien que l'effet d'interaction Contexte x Niveau soit non significatif. Rappelons que le traitement Contexte-1, qui fournit aux lecteurs un contexte visuel minimal d'un mot, produit significativement moins d'erreurs que le traitement sans contexte (Contexte-O) mais aussi moins d'erreurs que les traitements qui présentent un contexte plus élaboré (Contexte-2, Contexte-3 et Contexte-4).

Ceci dit, on constate néanmoins que, lorsqu'on restreint l'analyse à certains types de mots, les lecteurs profitent de la présence du contexte visuel en ce sens qu'ils commettent moins d'erreurs. Ainsi, en ce qui concerne l'article du syntagme nominal (tableau 3), les lecteurs de 3e année font significativement plus d'erreurs en l'absence du contexte qu'en sa présence (21.59 % vs 16.79 %). Et, lorsque l'analyse statistique porte sur le premier mot de la phrase (tableau 4), on constate que les lecteurs commettent significativement plus d'erreurs en absence du contexte (16.04 % vs. 13.02 %); nous y reviendrons plus loin.

L'étude du contexte visuel, dans une perspective de développement, indique que les lecteurs de 3e et 6e années se comportent différemment sur les plans du temps de lecture et du nombre d'erreurs. Ainsi, les lecteurs de 3e année lisent significativement moins vite en présence du contexte (3215 ms) qu'en son absence (3109 ms), tandis que cette différence est non significative pour les lecteurs de 6e année (2232 ms et

2230 ms). Des interactions analogues entre le contexte et le niveau scolaire se manifestent dans le cas du nombre d'erreurs. D'après les analyses de variance calculées sur le facteur Contexte à deux niveaux (présence vs. absence), l'effet d'interaction traduit le fait que les lecteurs de 3e année commettent significativement plus d'erreurs en absence du contexte qu'en sa présence (16,48 % vs 14,36 %, $Chi^2 = 10,47$, 1 ddl, $p < 0,001$), tandis que cette différence est non significative pour les lecteurs de 6e année (8,78% vs 7,84%).

Rappelons aussi la triple interaction entre les facteurs Syntagme, Contexte et Niveau (cf. tableau 3) qui indique que seuls les lecteurs de 3e année commettent significativement plus d'erreurs en absence du contexte dans le cas de l'article (21,59% vs 16,79%; $Chi^2 = 7,96$, 1 ddl, $p < 0,01$), contrairement aux lecteurs de 6e année (10,95 % vs. 10,63 %; $Chi^2 = 0.05$, 1ddl, $p = 0,818$).

Pour bien comprendre les implications psycholinguistiques de cette triple interaction, il est utile de se rappeler qu'en l'absence du contexte visuel suivant, le lecteur doit sélectionner l'article du syntagme nominal sans voir le nom qui, pourtant, exerce une contrainte syntaxique cruciale sur l'article (cf. Figure 2, version Contexte-0). Par ailleurs, on constate que la sélection de l'article pose plus de problèmes (temps de lecture plus long, nombre d'erreurs plus élevé) que celle du nom. Or, l'article fait partie d'une catégorie grammaticale, quelquefois appelée mots-outils, qui n'est pas porteuse d'un sens élaboré. Le temps de lecture prolongé et le pourcentage d'erreurs élevé ne semblent donc pas refléter une difficulté reliée au traitement sémantique de l'article lui-même; ils suggèrent plutôt que l'article représente une sorte de charnière à partir de laquelle commence à opérer un processus d'anticipation visant à sélectionner le ou les mots suivants, y compris le nom du syntagme nominal. Le fait que les enfants de 6e année ne commettent pas moins d'erreurs en lien avec l'article en présence du contexte visuel, pourrait alors signifier qu'ils ont développé des processus plus efficaces pour compenser l'absence visuelle du mot suivant: soit en regard du contexte précédent visuel (accès aux informations textuelles affichées à l'écran), soit en regard du contexte précédent non visuel (réactivation des informations textuelles emmagasinées).

De façon plus générale, l'étude de l'effet du contexte visuel illustre bien que l'existence d'une deuxième mesure dépendante, à côté de celle du temps de lecture, nous procure un renseignement complémentaire indispensable. En effet, si le temps de lecture était le seul indicateur pour se prononcer sur le phénomène en question,

on serait forcé d'admettre que l'absence du contexte exerce un effet facilitateur, une conclusion clairement contre-intuitive. Dans le cas de la technique Zigzag, le nombre d'erreurs fournit le renseignement qui permet de relativiser la portée des temps de lecture mesurés.

Selon une hypothèse de travail que nous étudions actuellement, le temps de lecture obtenu dans une tâche Zigzag pourrait être un indice de la *profondeur sémantique* des processus de lecture engagés et le nombre d'erreurs, l'indice de la *réussite* dans cette entreprise. Vu sous cet angle, le lecteur, en l'absence du contexte visuel, traiterait l'information affichée à l'écran moins en profondeur (= moins de temps), ce qui se traduirait par une moins bonne compréhension du texte lu (= plus d'erreurs). Selon cette interprétation, le temps prolongé serait alors la cause directe de la diminution du nombre d'erreurs.

Comparaison de deux études menées à l'aide de la technique Zigzag

L'étude démontre que plusieurs variables relatives au matériel de lecture interagissent significativement avec l'âge des lecteurs (3e année vs. 6 année). Ces interactions permettent d'identifier les phénomènes qui évoluent avec l'âge. A l'inverse, les variables qui n'interagissent pas avec l'âge sont susceptibles de nous indiquer des processus cognitifs plus fondamentaux, moins assujettis au changement. Dans cet ordre d'idées, signalons que la position des mots au sein de la phrase (premier mot, dernier mot, autres mots) et du texte (premier tiers, deuxième tiers, troisième tiers) n'a pas interagi avec le niveau scolaire. Dans une autre étude récente menée à l'aide de la technique Zigzag (Reinwein, 1992), l'analyse des mots en fonction de leur catégorie grammaticale et de leur position dans la phrase montre des résultats analogues. Ce parallélisme des résultats est d'autant plus remarquable que les deux études portent sur deux populations nettement distinctes: l'une est composée d'écoliers québécois du primaire qui lisent dans leur langue maternelle; l'autre est composée d'étudiants arabophones de l'Université de Tunis 1, lisant des textes en français, langue seconde.

Comparaison de la technique Zigzag avec d'autres techniques APS

La technique Zigzag partage plusieurs caractéristiques avec les techniques d'auto-présentation segmentée (APS) déja existantes. Ainsi:

- Le texte à lire y est présenté comme une suite de segments textuels, sans possibilité de retour en arrière.
- Le lecteur exerce un contrôle direct sur la durée d'affichage de chaque segment.
- La durée d'affichage de chaque segment donne lieu à une mesure de temps.

Par ailleurs, malgré son appellation de technique APS, la technique Zigzag possède plusieurs caractéristiques qui lui sont propres:

- Le type d'affichage du texte nécessite de la part du lecteur une sélection lexicale continue, chaque mot du texte (mot cible) étant jumelé avec un mot distracteur.
- Les mesures dépendantes se rattachent généralement à une partie et non à l'ensemble du segment textuel affiché à l'écran (sauf en version Contexte-0).
- L'ampleur du contexte visuel qui précède ou suit le mot cible, est variable: le contexte visuel précédant le mot cible peut comprendre jusqu'à environ quinze mots (une ligne) et le contexte suivant, un maximum de quatre mots.
- Chaque choix lexical est identifié comme une réussite ou une erreur et sert de deuxième mesure dépendante.
- On dispose aussi d'un indice de corrélation entre le temps de lecture et le nombre d'erreurs, indice calculable soit par mots, soit par sujets. Dans le premier cas, le temps moyen de chaque mot du texte (réussites seulement) est corrélé avec le nombre total d'erreurs portant sur ce mot. Dans le deuxième cas, pour chaque lecteur, le temps moyen des mots réussis d'un texte est corrélé avec le nombre total d'erreurs qu'il commet dans ce texte.

Sans doute, la technique Zigzag paraît plus artificielle que les techniques APS plus anciennes. En effet, en situation de lecture normale, un lecteur n'a pas à sélectionner les mots cibles. De plus, la technique Zigzag entraîne une diminution considérable de la vitesse de lecture. Nous croyons cependant que c'est là le prix à payer pour obtenir une deuxième mesure en direct. Le principe d'une sélection lexicale continue entre le mot cible et le mot distracteur, qui est probablement la source principale du ralentissement de lecture, oblige le lecteur non seulement à poser un geste mesurable (on peut mesurer le laps de temps écoulé entre deux commandes successives) mais aussi un geste dont la réussite est vérifiable (le choix du mot cible permet de conclure qu'il y a compréhension, locale tout au moins).

Ceci dit, nous croyons que, au lieu de se perdre dans des conjectures sur ce qu'est la lecture normale et véritable, il faut valider la technique Zigzag par le biais de

comparaisons systématiques soit avec d'autres techniques APS, soit avec d'autres tests traditionnels de lecture et d'aptitude de lecture (choix multiple, closure, etc.). Plusieurs études ont déjà été entreprises à cet égard ou sont sur la voie de l'être. En ce qui concerne la comparaison de la technique Zigzag avec d'autres techniques d'observation en direct, au moins deux des variables de la présente étude (position des mots dans la phrase, catégorie grammaticale) se prêtent bien à une analyse comparative pour mieux comprendre les ressemblances et différences des techniques en question. Examinons-les de plus près.

Un des constats les plus stables à travers les différentes études "on line" a trait au temps de lecture prolongé en fin de phrase (voir les références citées en introduction). Ce fait est généralement interprété comme indice d'une charge cognitive élevée du lecteur; elle découlerait d'une activité d'intégration qui vise une compréhension globale de la proposition ou de la phrase lue. Or, dans la présente étude, le temps du dernier mot de la phrase est significativement plus court que celui des autres mots et le nombre d'erreurs, nettement inférieur. Non seulement la même tendance se manifeste-t-elle aux deux niveaux scolaires, mais en plus un résultat analogue a été obtenu dans une autre étude auprès d'étudiants lisant des textes en langue seconde (voir plus haut).

Comment l'expliquer? Malgré l'apparente incompatibilité des résultats, nous croyons que les deux types de techniques sont sensibles au même phénomène psycholinguistique, mais qu'ils le reflètent différemment. En effet, dans le cas de la technique Zigzag, la charge cognitive élevée attribuable au processus d'intégration de la phrase pourrait être associée - sur le plan technique - non pas au dernier mot de la phrase mais au mot suivant celui-ci, c'est-à-dire au premier mot de la phrase suivante: le tableau 4 montre que le temps de lecture et le nombre d'erreurs associés au premier mot de la phrase sont significativement plus élevés. Le temps de lecture prolongé et le nombre d'erreurs élevé au premier mot de la phrase pourraient indiquer que: (i) le lecteur éprouve une difficulté de traitement du mot lui-même; (ii) le lecteur essaie d'anticiper la phrase à venir; (iii) le lecteur essaie de comprendre la phrase précédente.

L'hypothèse (i) paraît peut probable. Le premier mot (cible) de la phrase commence toujours par une lettre majuscule et est jumelé dans une proportion très élevée avec un mot distracteur commençant par une lettre minuscule. Cet indice typographique, à lui seul, devrait grandement faciliter l'identification du premier mot (cible) de la

phrase, ce qui n'est pas le cas: les lecteurs ont de la difficulté à le sélectionner rapidement et correctement. On peut alors présumer que l'effort supplémentaire qui se traduit par les deux mesures n'est pas investi dans le mot lui-même mais qu'il sert plutôt au traitement des segments textuels qui le précèdent ou suivent.

L'hypothèse (ii) à l'effet que le lecteur anticipe la phrase à venir paraît séduisante, d'autant plus que le premier mot d'une phrase semble être le lieu tout indiqué pour le déclenchement d'un tel processus d'anticipation. A ce sujet, il n'est pas sans intérêt que l'effet d'interaction de l'analyse de variance à trois facteurs (position, contexte, niveau scolaire) est près du seuil de signification [$F(2,32681) = 2,43$; $p = 0,09$]. Cette interaction traduit le fait que la présence ou l'absence du contexte visuel affecte bien plus le premier mot (différence de 242 ms) que le dernier mot (différence de 82 ms) ou les autres mots de la phrase (différence de 51 ms). La deuxième mesure se comporte de façon analogue: en ce qui concerne le pourcentage d'erreurs, la différence est de 3.02 % dans le cas du premier mot de la phrase, mais seulement de 0.78 % et 1.48 % respectivement, dans le cas du dernier mot et des autres mots. Ainsi, la sensibilité particulière du premier mot de la phrase pourrait être interprétée comme indice que le temps prolongé en présence du contexte visuel ait servi - au moins en partie - à une activité prospective du lecteur: un lecteur préoccupé par l'intégration des informations du contexte *précédent* ne devrait pas être affecté par la présence ou l'absence du contexte visuel qui *suit*. Un lecteur préoccupé des mots à venir, par contre, le devrait.

L'hypothèse (iii) voulant que le lecteur soit engagé dans une activité rétrospective, paraît également plausible. Son attrait principal réside dans le fait qu'elle permet d'établir un lien entre le temps peu élevé du dernier mot de la phrase et le temps élevé du premier mot de la phrase suivante (effet de débordement) et d'ouvrir ainsi la voie à une explication compatible avec les temps prolongés obtenus dans plusieurs études à la fin de la phrase. En effet, on présume généralement qu'une partie du temps de lecture peut être investie dans des processus qui impliquent la compréhension d'unités sémantiques plus larges, et que ces processus soient déclenchés à la frontière d'unités linguistiques telles que la proposition ou la phrase. Or, une frontière se situe entre deux unités successives. La question est alors de savoir si cette activité "frontalière" a des répercussions sur le dernier mot de l'unité x ou plutôt sur le premier mot de l'unité x + 1 (à moins que ce soit les deux). Le temps de lecture peu élevé associé au dernier mot suggère qu'un effet de débordement se produise avec, comme résultat, le transfert d'une partie du temps au mot suivant.

Reste à savoir pourquoi cet effet de débordement du dernier mot de la phrase (Mitchell, 1984: "processing spillover") devrait se manifester uniquement dans le cas de la technique Zigzag, mais non dans celui des autres techniques APS. Dans les deux cas, le mécanisme de mesure du temps d'affichage d'un segment textuel est pourtant essentiellement le même: le temps de lecture du segment suivant débute dès que le lecteur, ayant traité cognitivement le segment précédent, appuie sur la ou les touches du clavier relié au moniteur.

La raison pourrait être la suivante. Dans le cas de la technique Zigzag, le lecteur appuie sur l'une des deux touches dès qu'il a résolu le problème local de sélection du mot cible. En faisant cela, le lecteur amorce le temps de lecture associé au mot suivant (premier mot de la phrase suivante), et ce n'est qu'à ce moment que le lecteur effectue le travail d'intégration. Dans le cas des autres techniques APS par contre, en l'absence d'une tâche locale clairement délimitée, le lecteur pourrait retarder son geste et n'appuyer sur la touche du clavier qu'*après* avoir terminé le traitement intégral de la phrase.

Comme corollaire de l'hypothèse de travail (iii), le temps de lecture peu élevé du dernier mot de la phrase serait le reflet direct de son traitement cognitif peu exigeant. En effet, au fur et à mesure que le lecteur approche de la fin d'une phrase, il devient de plus en plus facile de prédire le mot à venir.

Une mise en garde s'impose ici. Même si le dernier mot de la phrase est associé au temps le plus court et que le premier mot est associé au temps le plus long (constat analogue pour les erreurs), il ne faut pas s'attendre à ce qu'il y ait un lien linéaire quelconque entre ces deux mots[1]: il n'existe pas de corrélation significative entre le dernier et le premier mot de la phrase, ni pour le temps et le nombre d'erreurs (Reinwein, 1992). Ce résultat confirme notre intuition à ce sujet: si le premier mot de la phrase assume le temps supplémentaire nécesssaire au traitement d'intégration de la phrase *dans son ensemble* , on devrait s'attendre à ce que le premier mot soit corrélé avec des facteurs relatifs à la phrase dans son ensemble (ou d'une partie de

[1] Dans une autre étude (Reinwein, 1992), à partir de deux textes, nous avons vérifié les liens qui existent entre les mots qui se suivent immédiatement (r de Pearson). Ces corrélations ont révélé une cohésion interne élevée entre l'article et le nom qui suit [r = 0.80 (texte 1) et 0.64 (texte 2); p < 0,001], cohésion qui contraste singulièrement avec l'absence de lien entre l'article et le mot précédent [r = -0,35 (texte 1) et -0,35 (texte 2); ns] d'une part, et entre le nom et le mot suivant [r = -0,05 (texte 1) et -0,18 (texte 2); ns], d'autre part. En ce qui concerne les erreurs, des corrélations analogues, quoique moins prononcées, ont été obtenues.

celle-ci) plutôt qu'avec le dernier mot seulement. C'est une voie à explorer sur la base d'un échantillon textuel plus large.

Une deuxième différence majeure entre la technique Zigzag et les autres techniques APS porte sur les temps de lecture des mots selon leur catégorie grammaticale et selon leur longueur. Ces deux variables ont d'ailleurs été examinées à plusieurs reprises dans des études en direct (voir les références citées en introduction). Généralement, les mots courts donnent lieu à des temps de lecture moins élevés. D'après Haberlandt (1984), cet effet est général au point d'avoir été obtenu avec différentes méthodes d'auto-présentation. Or, dans la présente étude, l'effet de longueur des mots nous indique une tendance diamétralement opposée à ces résultats: selon la technique Zigzag, les temps de lecture des mots plus courts (1 à 4 lettres) sont supérieurs à ceux des mots plus longs (5 à 9 lettres). De façon analogue, la dichotomie entre mots-outils et mots pleins, introduite comme variable indépendante dans plusieurs études d'observation en direct, révèle la même incompatibilité avec les résultats recueillis selon la technique Zigzag: les catégories grammaticales correspondant aux mots-outils (pronom, préposition, conjonction, article) prennent plus, et non pas moins, de temps que les mots pleins (nom et verbe en particulier). Rappelons à ce sujet l'effet d'interaction entre ces deux variables qui s'explique par le fait que les mots-outils sont généralement des mots courts et les mots pleins, des mots longs.

L'hypothèse de travail pour expliquer cette apparente incompatibilité des techniques d'observation nous paraît de nouveau l'hypothèse du débordement d'un mot à l'autre: dans le cas de la technique Zigzag, une partie du temps de traitement (celle consacrée au traitement d'unités sémantiques plus larges) serait mesurée en décalage, comparativement à d'autres techniques d'observation en direct. Cette hypothèse ne se veut cependant pas exclusive. Des facteurs perceptifs pourraient également jouer un rôle important: le type d'affichage des mots selon la technique Zigzag pourrait neutraliser une partie de l'effet attribuable à la longueur des mots, tel qu'identifié dans certaines études d'observation des mouvements oculaires.

En résumé, il est évident que la technique Zigzag se distingue de façon importante des autres techniques d'observation en direct. Seule une comparaison directe des diverses techniques APS, dans des conditions expérimentales contrôlées (même matériel de lecture, mêmes sujets, etc.) permettra des réponses plus définitives. Ce travail reste encore à faire.

Remerciements

Je tiens à remercier tout particulièrement Lucie Huberdeau pour le traitement statistique des données et ses commentaires pertinents, Serge Tassé pour l'expérimentation dans les classes de troisième année et Andrée Duchesneau pour la révision du texte.

Bibliographie

Aaronson, D. et Ferres, S.: A quantitative model for coding lexical categories. Journal of Experimental Psychology: Human Perception and Performance. 1983, 9, 700-725.

Aaronson, D. et Ferres, S.: The word-by-word reading paradigm: An experimental and theoretical approach. In: D. E. Kieras et M. A. Just (Eds.): New Methods in Reading Comprehension Research. Hillsdale, New Jersey: Lawrence Erlbaum, 1984a, 31-68.

Aaronson, D. et Ferres, S.: A structure and meaning based classification of lexical categories. In: S. White, B. Kachuck et F. Podwall (Eds.): Reading and Reading Disability: Scanning the 80's. New York, The N. Y. Academy of Sciences, 1984b.

Booher, H. R. (1975): "Relative comprehensibility of pictorial information and printed words in proceduralized instructions". Human Factors, 17 (3), 266-277.

Ciesielski, R. et Reinwein, J.: Le test Zigzag (version 2.0, LightSpeed Pascal). Université du Québec à Montréal, Département de linguistique, 1989.

Cirilo, R. K. et Foss, D. J.: Text structure and reading time for sentences. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 96-109.

Ehrlich, K. et Rayner, K.: Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. Journal of Verbal Learning and Verbal Behavior, 1983, 22, 75-82.

Ferres, S. et Aaronson, D.: A word class encoding model for reading units. The Psychonomic Society, 1981.

Frochot, M., Zagar, D. et Fayol, M.: Effets de l'organisation narrative sur la lecture de récits Année Psychologique, 1987, 87, 237-252.

Graesser, A. C., Hoffman, N. L. et Clark, L. F.: Structural components of reading time. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 135-151.

Green, D. W., Mitchell, D. C. et Hammond, E. J.: The scheduling of text integration processing in reading. Quarterly Journal of Experimental Psychology, 1981, 33A, 455-464.

Guttentag, R. E. et Haith, M. M.: A longitudinal study of word processing by first-grade children. Journal of Educational Psychology, 1980, 72, 701-705.

Haberlandt, K.: Components of sentence and word reading times. In: D. E. Kieraset M. A. Just, (Eds.): New Methods in Reading Comprehension Research. Hillsdale, New Jersey: Lawrence Erlbaum, 1984, 219-252.

Haberlandt, K. et Graesser, A. C.: Component processes in text comprehension and some of their interactions. Journal of Experimental Psychology: General, 1985, 3, 357-374.

Haberlandt, K., Berian, C. et Sandson, J.: The episode schema in story processing. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 635-650.

Jarry, C.: Le test Zigzag: comparaison avec un questionnaire à choix multiples et le rendement scolaire en lecture. Université du Québec à Montréal, rapport du projet d'intervention (maîtrise en enseignement au primaire), 1993.

Just, M. A. et Carpenter, P. A.: A theory of reading: From eye fixations to comprehension. Psychological Review, 1980, 87(4), 329-354.

Just, A. M., Carpenter, P. A. et Woolley, J. D.: Paradigms and processes in reading comprehension. Journal of Experimental Psychology: General, 1982, 111(2), 228-238.

Levy-Schoen, Ariane: Les mouvements des yeux comme indicateurs des processus cognitifs. In: J.-P. Caverni, C. Bastien, P. Mendelsohn et G. Tiberghien (Eds.): Psychologie cognitive: modèles et méthodes. Grenoble: Presses universitaires de Grenoble, 1988, 329-347.

Mitchell, D. C. et Green, D. W.: The effects of context and content on immediate processing in reading. Quarterly Journal of Experimental Psychology, 1978, 30, 609-636.

Mitchell, D. C. et Holmes, V. M.: The role of specific information about the verb in parsing sentences with local structural ambiguity. Journal of Memory and Language, 1985, 24, 542-559.

Pynte, Joël: Lire ... Identifier, comprendre. Presses universitaires de Lille, 1983.

Reinwein, J.: La technique Zigzag comme outil pour mesurer l'effet de l'illustration et du texte sur le lecteur en langue seconde. In: C. Préfontaine et M. Lebrun (Eds.): Stratégie d'enseignement et d'apprentissage en lecture / écriture: actes de colloque. Montréal: Editions Logiques, 1992, 261-306.

Rossi, J-P.: Les mécanismes de la lecture. Paris: Publications de la Sorbonne, 1985.

Sternberg, S.: Memory-scanning: Mental processes revealed by reaction-time experiments. American Scientist, 1969, 57, 421-457.

Zagar, D.: L'utilisation du temps d'exposition comme indicateur du temps de traitement pendant la lecture. In: J.-P. Caverni, C. Bastien, P. Mendelsohn et G. Tiberghien (Eds.): Psychologie cognitive: modèles et méthodes. Grenoble: Presses universitaires de Grenoble, 1988.

Zola, D.: Redundancy and word perception during reading. Perception & Psychophysics, 1984, 36, 277-284.

Ecran 1

Mon	étés	faire	mère,	juillet
les	père,	ma	pêche.	mon

Ecran 2

Mon

étés	faire	mère,	juillet	frère
père,	ma	pêche.	mon	pour

Ecran 3

Mon père,

faire	mère,	juillet	frère	et
ma	pêche.	mon	pour	du

Ecran 4

Mon père, ma

mère,	juillet	frère	et	moi,
pêche.	mon	pour	du	tant

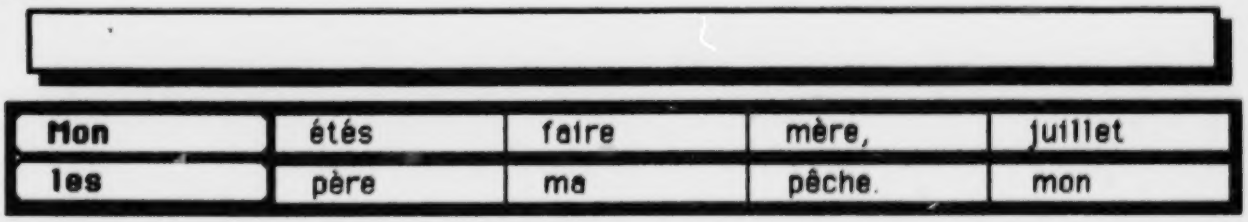
Ecran 5

Mon père, ma mère,

juillet	frère	et	moi,	raison
mon	pour	du	tant	nous

Figure 1

Contexte 4

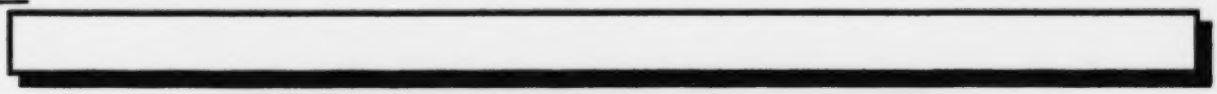


Mon	étés	faire	mère,	juillet
les	père	ma	pêche.	mon

Contexte 3

Mon	étés	faire	mère,
les	père	ma	pêche.

Contexte 2



Mon	étés	faire
les	père	ma

Contexte 1

Mon	étés
les	père

Contexte 0

Mon
les

Figure 2